

日 本 国 特 許 庁

PATENT OFFICE  
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されて  
る事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed  
with this Office.

出 願 年 月 日  
Date of Application:

1999年 9月10日

願 番 号  
Application Number:

平成11年特許願第257429号

願 人  
Applicant(s):

インターナショナル・ビジネス・マシーンズ・コーポレイシ  
ョン

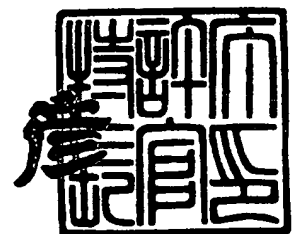


CERTIFIED COPY OF  
PRIORITY DOCUMENT

1999年11月12日

特許庁長官  
Commissioner,  
Patent Office

近 藤 隆 彦



出証番号 出証特平11-3079005

【書類名】 特許願

【整理番号】 JA999202

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 13/00

【発明者】

    【住所又は居所】 神奈川県大和市下鶴間 1 6 2 3 番地 1 4 日本アイ・ビー・エム株式会社 大和事業所内

    【氏名】 田原 義則

【発明者】

    【住所又は居所】 神奈川県大和市下鶴間 1 6 2 3 番地 1 4 日本アイ・ビー・エム株式会社 大和事業所内

    【氏名】 市川 統一

【発明者】

    【住所又は居所】 神奈川県大和市下鶴間 1 6 2 3 番地 1 4 日本アイ・ビー・エム株式会社 大和事業所内

    【氏名】 宇田 哲也

【発明者】

    【住所又は居所】 神奈川県大和市下鶴間 1 6 2 3 番地 1 4 日本アイ・ビー・エム株式会社 大和事業所内

    【氏名】 友田 大輔

【特許出願人】

    【識別番号】 390009531

    【住所又は居所】 アメリカ合衆国 1 0 5 0 4、ニューヨーク州アーモンク (番地なし)

    【氏名又は名称】 インターナショナル・ビジネス・マシーンズ・コーポレーション

【代理人】

    【識別番号】 100086243

    【弁理士】

【氏名又は名称】 坂口 博

【連絡先】 0 4 6 2 - 7 3 - 3 3 1 8、3 3 2 5、3 4 5 5

【選任した代理人】

【識別番号】 100091568

【弁理士】

【氏名又は名称】 市位 嘉宏

【手数料の表示】

【予納台帳番号】 024154

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9304391

【包括委任状番号】 9304392

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 認識単語登録方法、音声認識方法、音声認識装置、認識単語登録のためのソフトウェア・プロダクトを格納した記憶媒体、音声認識のためのソフトウェア・プロダクトを格納した記憶媒体

【特許請求の範囲】

【請求項 1】

表示画面、音声入力装置を備える音声認識装置上で実施される認識単語の登録方法であって、

- (a) ユーザによって特定された単語表記を取得する段階と、
  - (b) 単語辞書を検索することにより、前記単語表記に対応する読みを取得する段階と、
  - (c) 発音辞書を検索することにより、該取得した読みに対応するベースフォームを取得する段階と、
  - (d) 該取得したベースフォームを音声認識辞書に登録する段階と、
- を含む認識単語登録方法。

【請求項 2】

表示画面、音声入力装置を備える音声認識装置上で実施される認識単語の登録方法であって、

- (a) ユーザによって特定された単語表記を取得する段階と、
- (b) 単語辞書を検索することにより、前記単語表記に対応する複数の読み候補と各々が前記複数の読み候補に対応する読みスコアを取得する段階と、
- (c) ユーザにより前記複数の読み候補を表示する段階と、
- (d) ユーザにより前記複数の読み候補のうちの 1 つの読みが選択されたことをに応答して、該選択された読みを取得する段階と、
- (e) 発音辞書を検索することにより、前記選択された読みに対応するベースフォームと発音スコアを取得する段階と、
- (f) 前記発音スコアが所定の閾値を越えているか否かを判断する段階と、
- (g) 前記発音スコアが所定の閾値を越えている場合には、該取得したベースフォームを音声認識辞書に登録する段階と、

を含む認識単語登録方法。

【請求項 3】

表示画面、音声入力装置を備える音声認識装置上で実施される音声認識方法であって、

(a) ユーザの発声によって入力された第 1 の音声情報が所定の条件に一致するか否かを判断する段階と、

(b) 前記音声情報が所定の条件に一致すると判断した場合には、前記表示画面上に認識単語入力フィールドと、読み入力フィールドを有する音声認識ウィザード・パネルを表示する段階と、

(c) 前記音声認識ウィザード・パネルに入力された認識単語と読みを取得する段階と、

(d) 前記ユーザが、前記認識単語と前記読みに対応して行なった発声に基づく第 2 の音声情報を取得する段階と、

(e) 前記第 2 の音声情報と、前記認識単語と、前記読みとに基づいて、ベースフォームを特定する段階と、

(f) 前記ベースフォームを音声認識辞書に追加する段階と、  
を含む音声認識方法。

【請求項 4】

表示画面、音声入力装置を備える音声認識装置上で実施される音声認識方法であって、

(a 1) ユーザによって特定された単語表記を取得する段階と、

(a 2) 単語辞書を検索することにより、前記単語表記に対応する複数の読み候補と各々が前記複数の読み候補に対応する読みスコアを取得する段階と、

(a 3) ユーザに前記複数の読み候補を表示する段階と、

(a 4) ユーザにより前記複数の読み候補のうちの 1 つの読みが選択されたことをに応答して、該選択された読みを取得する段階と、

(a 5) 発音辞書を検索することにより、前記選択された読みに対応するベースフォームと発音スコアを取得する段階と、

(a 6) 前記発音スコアが所定の閾値を越えているか否かを判断する段階と、

(a 7) 前記発音スコアが所定の閾値を越えている場合には、該取得したベースフォームを音声認識辞書に登録する段階と、

を含む初期登録段階と、

(b 1) ユーザの発声によって入力された第 1 の音声情報が所定の条件に一致するか否かを判断する段階と、

(b 2) 前記音声情報が所定の条件に一致すると判断した場合には、前記表示画面上に認識単語入力フィールドと、読み入力フィールドを有する音声認識ウィザード・パネルを表示する段階と、

(b 3) 前記音声認識ウィザード・パネルに入力された認識単語と読みを取得する段階と、

(b 4) 前記ユーザが、前記認識単語と前記読みに対応して行なった発声に基づく第 2 の音声情報を取得する段階と、

(b 5) 前記第 2 の音声情報と、前記認識単語と、前記読みとに基づいて、第 2 のベースフォームを特定する段階と、

(b 6) 前記第 2 のベースフォームを音声認識辞書に追加する段階と、

を含む音声認識時登録段階と、

を含む音声認識方法。

#### 【請求項 5】

表示画面、音声入力装置を備える音声認識装置であって、

(a) ユーザによって特定された単語表記を取得する認識単語登録部と、

(b) 単語辞書を検索することにより、前記単語表記に対応する読みを取得する読み生成部と、

(c) 発音辞書を検索することにより、該取得した読みに対応するベースフォームを取得するベースフォーム生成部と、

(d) 該取得したベースフォームに登録する音声認識辞書と、

を含む音声認識装置。

#### 【請求項 6】

表示画面、音声入力装置を備える音声認識装置であって、

(a) ユーザによって特定された単語表記を取得する認識単語登録部と、

(b) 単語辞書を検索することにより、前記単語表記に対応する複数の読み候補と各々が前記複数の読み候補に対応する読みスコアを取得するとともに、

(c) 前記表示装置に表示した複数の読み候補のうちの1つの読みが選択されたことをに応答して、該選択された読みも取得する読み生成部と、

(d) 発音辞書を検索することにより、前記選択された読みに対応するベースフォームと発音スコアを取得するベースフォーム生成部と、

(e) 前記発音スコアが所定の閾値を越えている場合には、該取得したベースフォームが登録される音声認識辞書と、

を含む音声認識装置。

【請求項7】

表示画面、音声入力装置を備える音声認識装置であって、

(a) ユーザの発声によって入力された第1の音声情報が所定の条件に一致するか否かを判断する認識単語登録部と、

(b) 前記音声情報が所定の条件に一致すると判断した場合には、前記表示画面上に認識単語入力フィールドと、読み入力フィールドを有する音声認識ウィザード・パネルを表示する音声認識ウィザードと、

(c) 前記音声認識ウィザード・パネルに入力された認識単語と読みに対応して行なった発声に基づく第2の音声情報を取得する音声入力部と、

(d) 前記第2の音声情報と、前記認識単語と、前記読みとに基づいて、ベースフォームを特定するベースフォーム生成部と、

(f) 前記ベースフォームを追加登録する音声認識辞書と、  
を含む音声認識装置。

【請求項8】

表示画面、音声入力装置を備える音声認識装置であって、

(a) ユーザの発声により生じた音声情報を入力する音声入力部と、

(a) 前記音声情報を認識する音声認識エンジンと、

(b) ユーザによって特定された単語表記を取得する認識単語登録部と、

(c) 単語辞書を検索することにより、前記単語表記に対応する複数の読み候補と各々が前記複数の読み候補に対応する読みスコアを取得するとともに、ユー

ザにより前記複数の読み候補のうちの1つの読みが選択されたことをに応答して、該選択された読みを取得する読み生成部と、

(d) 発音辞書を検索することにより、前記選択された読みに対応するベースフォームと発音スコアを取得するベースフォーム生成部と、

(e) 前記発音スコアが所定の閾値を越えている場合には、該取得したベースフォームが登録される音声認識辞書と、

を含み、

(a2) 前記音声認識エンジンは、ユーザの発声によって入力された第1の音声情報が所定の条件に一致するか否かを判断し、前記音声情報が所定の条件に一致すると判断した場合には、前記表示画面上に認識単語入力フィールドと、読み入力フィールドを有する音声認識ウィザード・パネルを管理する音声認識ウィザードを起動し、

(c2) 前記読み生成部は、前記音声認識ウィザード・パネルに入力された認識単語と読みに対応して行なった発声に基づく第2の音声情報を取得し、

(d2) 前記ベースフォーム生成部は、前記第2の音声情報と、前記認識単語と、前記読みとに基づいて、第2のベースフォームを特定し、

(e2) 前記音声認識辞書は、前記第2のベースフォームを保持することを特徴とする音声認識装置。

#### 【請求項9】

表示画面、音声入力装置を備える音声認識装置上で実施される認識単語の登録処理プログラムを格納した記録媒体であって、

該プログラムは、

(a) ユーザによって特定された単語表記を取得することを前記音声認識装置に指示するためのプログラムコードと、

(b) 単語辞書を検索することにより、前記単語表記に対応する読みを取得することを前記音声認識装置に指示するためのプログラムコードと、

(c) 発音辞書を検索することにより、該取得した読みに対応するベースフォームを取得することを前記音声認識装置に指示するためのプログラムコードと、

(d) 該取得したベースフォームを音声認識辞書に登録することを前記音声認



識装置に指示するためのプログラムコードと、  
を含む記憶媒体。

【請求項 10】

表示画面、音声入力装置を備える音声認識装置上で実施される認識単語の登録  
処理プログラムを格納した記録媒体であって、

該プログラムは、

(a) ユーザによって特定された単語表記を取得することを前記音声認識装置  
に指示するためのプログラムコードと、

(b) 単語辞書を検索することにより、前記単語表記に対応する複数の読み候  
補と各々が前記複数の読み候補に対応する読みスコアを取得することを前記音声  
認識装置に指示するためのプログラムコードと、

(c) ユーザにより前記複数の読み候補を表示することを前記音声認識装置に  
指示するためのプログラムコードと、

(d) ユーザにより前記複数の読み候補のうちの 1 つの読みが選択されたこと  
をに応答して、該選択された読みを取得することを前記音声認識装置に指示する  
ためのプログラムコードと、

(e) 発音辞書を検索することにより、前記選択された読みに対応するベース  
フォームと発音スコアを取得することを前記音声認識装置に指示するためのプロ  
グラムコードと、

(f) 前記発音スコアが所定の閾値を越えているか否かを判断することを前記  
音声認識装置に指示するためのプログラムコードと、

(g) 前記発音スコアが所定の閾値を越えている場合には、該取得したベース  
フォームを音声認識辞書に登録することを前記音声認識装置に指示するためのプ  
ログラムコードと、

を含む記憶媒体。

【請求項 11】

表示画面、音声入力装置を備える音声認識装置上で実施される音声認識処理プ  
ログラムを格納した記録媒体であって、

該プログラムは、

(a) ユーザの発声によって入力された第1の音声情報が所定の条件に一致するか否かを判断することを前記音声認識装置に指示するためのプログラムコードと、

(b) 前記音声情報が所定の条件に一致すると判断した場合には、前記表示画面上に認識単語入力フィールドと、読み入力フィールドを有する音声認識ウィザード・パネルを表示することを前記音声認識装置に指示するためのプログラムコードと、

(c) 前記音声認識ウィザード・パネルに入力された認識単語と読みを取得することを前記音声認識装置に指示するためのプログラムコードと、

(d) 前記ユーザが、前記認識単語と前記読みに対応して行なった発声に基づく第2の音声情報を取得することを前記音声認識装置に指示するためのプログラムコードと、

(e) 前記第2の音声情報と、前記認識単語と、前記読みとに基づいて、ベースフォームを特定することを前記音声認識装置に指示するためのプログラムコードと、

(f) 前記ベースフォームを音声認識辞書に追加することを前記音声認識装置に指示するためのプログラムコードと、

を含む記憶媒体。

#### 【請求項12】

表示画面、音声入力装置を備える音声認識装置上で実施される音声認識処理プログラムを格納した記録媒体であって、

該プログラムは、

(a) ユーザによって特定された単語表記を取得することを前記音声認識装置に指示するためのプログラムコードと、

(b) 単語辞書を検索することにより、前記単語表記に対応する複数の読み候補と各々が前記複数の読み候補に対応する読みスコアを取得することを前記音声認識装置に指示するためのプログラムコードと、

(c) ユーザに前記複数の読み候補を表示することを前記音声認識装置に指示するためのプログラムコードと、

(d) ユーザにより前記複数の読み候補のうちの1つの読みが選択されたことをに応答して、該選択された読みを取得することを前記音声認識装置に指示するためのプログラムコードと、

(e) 発音辞書を検索することにより、前記選択された読みに対応するベースフォームと発音スコアを取得することを前記音声認識装置に指示するためのプログラムコードと、

(f) 前記発音スコアが所定の閾値を越えているか否かを判断することを前記音声認識装置に指示するためのプログラムコードと、

(g) 前記発音スコアが所定の閾値を越えている場合には、該取得したベースフォームを音声認識辞書に登録することを前記音声認識装置に指示するためのプログラムコードと、

(h) ユーザの発声によって入力された第1の音声情報が所定の条件に一致するか否かを判断することを前記音声認識装置に指示するためのプログラムコードと、

(i) 前記音声情報が所定の条件に一致すると判断した場合には、前記表示画面上に認識単語入力フィールドと、読み入力フィールドを有する音声認識ウィザード・パネルを表示することを前記音声認識装置に指示するためのプログラムコードと、

(j) 前記音声認識ウィザード・パネルに入力された認識単語と読みを取得することを前記音声認識装置に指示するためのプログラムコードと、

(k) 前記ユーザが、前記認識単語と前記読みに対応して行なった発声に基づく第2の音声情報を取得することを前記音声認識装置に指示するためのプログラムコードと、

(l) 前記第2の音声情報と、前記認識単語と、前記読みとに基づいて、第2のベースフォームを特定することを前記音声認識装置に指示するためのプログラムコードと、

(m) 前記第2のベースフォームを音声認識辞書に追加することを前記音声認識装置に指示するためのプログラムコードと、

を含む記憶媒体。

【発明の詳細な説明】

【0001】

【産業上の利用分野】

この発明は、音声認識辞書のカスタマイズする方式に関するものであり、より詳しくは、ユーザが行なう音声認識辞書への登録手順が軽減された音声認識辞書のカスタマイズ方式に関する。

【0002】

【従来の技術】

従来より、日本語等の音声認識において、ユーザが認識可能な単語を登録するに際しては、認識すべき単語表記（漢字・英字等を含む）に、

1. その読みを入力し、
2. そのよみに対して、予測される複数のベースフォーム（発音）を、ユーザの発声した発音と比較し、最も評価値が高く、かつ所定の閾値を越えるものを採用し、音声認識辞書に登録していた。

【0003】

かかるユーザの登録作業を軽減するために、読みの入力に関しては、かな漢入力時のキーストロークを得ることにより、よみの入力の手間をはぶくことも考えられる。しかし、かな漢を使用しない場合や、キーストロークからよみを得られない場合、英単語などの場合にはこの手法は利用できないという問題がある。

【0004】

また、日本語等では、よみ（かな）は発音に一対一で対応していないことが多く、実際の発声による情報によりベースフォームを選択しなければ、音声認識精度を高いレベルで維持することができない。例えば、「は」、「へ」、「を」や長音化（ー）は単語によって大きな揺らぎが発生する。単純化して説明すれば、「こんにちは」は"ko\_nnichiw a"である。また、その外の例として「子牛(こうし)」は"kou sh i"であるが「講師(こうし)」は"ko o: sh i"など同一の読みでも別のベースフォームを使用しなければならないものが存在する。

【0005】

このため、図10、11に示すように、従来の音声認識ソフトウェアにおいて

は、登録する単語 501 を特定し、その読み、発音表記を入力パネル 500 のフィールド 507、509 に入力し、録音ボタン 503 を押して、実際に発音することにより、音声認識を行なう単語の音声認識辞書への登録を行なっている。

【0006】

そして、入力された音声情報と、その読みに対応する複数のベースフォームを順次比較し、最も評価値が高いものが所定の閾値を越えているか否かを検査する。最も評価値の高いベースフォームが所定の閾値を越えている場合には、そのベースフォームを単語表記 501、読み 507、発音表記 509 とともに、音声認識辞書に登録する。

【0007】

この一方、最も評価値の高いベースフォームが所定の閾値を越えていない場合には、再度ユーザに発声を依頼するパネル 520 を表示し、入力された音声情報を基に、再度検査を行なう。この作業は、評価値が所定の閾値を越えるまで繰り返され、ユーザの負担となっていた。

【0008】

【発明が解決しようとする課題】

本発明の目的は、認識対象になっていない単語を音声認識辞書に登録するに際し、音声認識の精度を維持しつつ、ユーザーの音声による録音の作業を可能な限り低減させることのできる音声認識システムを提供することにある。

【0009】

本発明の他の一つの目的は、登録するユーザの音声情報を可能な限り減少させることにより、必要となる資源の低減された音声認識システムを提供することにある。

【0010】

本発明の他の一つの目的は、ユーザが直感的に把握できる使い勝手のよい音声認識システムを提供することにある。

【0011】

【課題を解決するための手段】

単語表記から読み群をスコア付きで生成する。生成された読み群からユーザー

が正しいよみを選択する。選択された読みからと単語表記-よみからベースフォーム群をスコア付きで生成し、スコアが基準値を超えるものを声の録音なしで登録する。ユーザが音声認識処理を行なうに際し、認識エラーがN回を超えた場合等所定のエラー状態を検出した場合には、ユーザに単語表記、読み、発音表記を登録し、対応する音声を発声することを促すパネルを出力する。ユーザーの実際の発声にもとづいてベースフォームを取得し、音声認識辞書に再登録する。

【0012】

本発明の一態様においては、

表示画面、音声入力装置を備える音声認識装置上で実施される認識単語の登録方法であって、

- (a) ユーザによって特定された単語表記を取得する段階と、
  - (b) 単語辞書を検索することにより、前記単語表記に対応する読みを取得する段階と、
  - (c) 発音辞書を検索することにより、該取得した読みに対応するベースフォームを取得する段階と、
  - (d) 該取得したベースフォームを音声認識辞書に登録する段階と、
- を含む認識単語登録方法が提供される。

【0013】

本発明の他の一態様においては、

表示画面、音声入力装置を備える音声認識装置上で実施される認識単語の登録方法であって、

- (a) ユーザによって特定された単語表記を取得する段階と、
- (b) 単語辞書を検索することにより、前記単語表記に対応する複数の読み候補と各々が前記複数の読み候補に対応する読みスコアを取得する段階と、
- (c) ユーザにより前記複数の読み候補を表示する段階と、
- (d) ユーザにより前記複数の読み候補のうちの1つの読みが選択されたことをに応答して、該選択された読みを取得する段階と、
- (e) 発音辞書を検索することにより、前記選択された読みに対応するベースフォームと発音スコアを取得する段階と、

(f) 前記発音スコアが所定の閾値を越えているか否かを判断する段階と、  
(g) 前記発音スコアが所定の閾値を越えている場合には、該取得したベースフォームを音声認識辞書に登録する段階と、  
を含む認識単語登録方法が提供される。

【0014】

本発明の他の一態様においては、  
表示画面、音声入力装置を備える音声認識装置上で実施される音声認識方法であって、

(a) ユーザの発声によって入力された第1の音声情報が所定の条件に一致するか否かを判断する段階と、

(b) 前記音声情報が所定の条件に一致すると判断した場合には、前記表示画面上に認識単語入力フィールドと、読み入力フィールドを有する音声認識ウィザード・パネルを表示する段階と、

(c) 前記音声認識ウィザード・パネルに入力された認識単語と読みを取得する段階と、

(d) 前記ユーザが、前記認識単語と前記読みに対応して行なった発声に基づく第2の音声情報を取得する段階と、

(e) 前記第2の音声情報と、前記認識単語と、前記読みとに基づいて、ベースフォームを特定する段階と、

(f) 前記ベースフォームを音声認識辞書に追加する段階と、  
を含む音声認識方法が提供される。

【0015】

本発明の他の一態様においては、  
表示画面、音声入力装置を備える音声認識装置上で実施される音声認識方法であって、

(a1) ユーザによって特定された単語表記を取得する段階と、

(a2) 単語辞書を検索することにより、前記単語表記に対応する複数の読み候補と各々が前記複数の読み候補に対応する読みスコアを取得する段階と、

(a3) ユーザに前記複数の読み候補を表示する段階と、

(a 4) ユーザにより前記複数の読み候補のうちの 1 つの読みが選択されたことをに応答して、該選択された読みを取得する段階と、

(a 5) 発音辞書を検索することにより、前記選択された読みに対応するベースフォームと発音スコアを取得する段階と、

(a 6) 前記発音スコアが所定の閾値を越えているか否かを判断する段階と、

(a 7) 前記発音スコアが所定の閾値を越えている場合には、該取得したベースフォームを音声認識辞書に登録する段階と、

を含む初期登録段階と、

(b 1) ユーザの発声によって入力された第 1 の音声情報が所定の条件に一致するか否かを判断する段階と、

(b 2) 前記音声情報が所定の条件に一致すると判断した場合には、前記表示画面上に認識単語入力フィールドと、読み入力フィールドを有する音声認識ウィザード・パネルを表示する段階と、

(b 3) 前記音声認識ウィザード・パネルに入力された認識単語と読みを取得する段階と、

(b 4) 前記ユーザが、前記認識単語と前記読みに対応して行なった発声に基づく第 2 の音声情報を取得する段階と、

(b 5) 前記第 2 の音声情報と、前記認識単語と、前記読みとに基づいて、第 2 のベースフォームを特定する段階と、

(b 6) 前記第 2 のベースフォームを音声認識辞書に追加する段階と、

を含む音声認識時登録段階と、

を含む音声認識方法が提供される。

【0 0 1 6】

本発明の他の一態様においては、

表示画面、音声入力装置を備える音声認識装置であって、

(a) ユーザによって特定された単語表記を取得する認識単語登録部と、

(b) 単語辞書を検索することにより、前記単語表記に対応する読みを取得する読み生成部と、

(c) 発音辞書を検索することにより、該取得した読みに対応するベースフォ



ームを取得するベースフォーム生成部と、

(d) 該取得したベースフォームを登録する音声認識辞書と、  
を含む音声認識装置が提供される。

【0017】

本発明の他の一態様においては、

表示画面、音声入力装置を備える音声認識装置であって、

(a) ユーザによって特定された単語表記を取得する認識単語登録部と、

(b) 単語辞書を検索することにより、前記単語表記に対応する複数の読み候補と各々が前記複数の読み候補に対応する読みスコアを取得するとともに、

(c) 前記表示装置に表示した複数の読み候補のうちの1つの読みが選択されたことをに応答して、該選択された読みも取得する読み生成部と、

(d) 発音辞書を検索することにより、前記選択された読みに対応するベースフォームと発音スコアを取得するベースフォーム生成部と、

(e) 前記発音スコアが所定の閾値を越えている場合には、該取得したベースフォームが登録される音声認識辞書と、

を含む音声認識装置が提供される。

【0018】

本発明の他の一態様においては、

表示画面、音声入力装置を備える音声認識装置であって、

(a) ユーザの発声によって入力された第1の音声情報が所定の条件に一致するか否かを判断する認識単語登録部と、

(b) 前記音声情報が所定の条件に一致すると判断した場合には、前記表示画面上に認識単語入力フィールドと、読み入力フィールドを有する音声認識ウィザード・パネルを表示する音声認識ウィザードと、

(c) 前記音声認識ウィザード・パネルに入力された認識単語と読みに対応して行なった発声に基づく第2の音声情報を取得する音声入力部と、

(d) 前記第2の音声情報と、前記認識単語と、前記読みとに基づいて、ベースフォームを特定するベースフォーム生成部と、

(f) 前記ベースフォームを追加登録する音声認識辞書と、

を含む音声認識装置が提供される。

【0 0 1 9】

本発明の他の一態様においては、

表示画面、音声入力装置を備える音声認識装置であって、

(a) ユーザの発声により生じた音声情報を入力する音声入力部と、

(a) 前記音声情報を認識する音声認識エンジンと、

(b) ユーザによって特定された単語表記を取得する認識単語登録部と、

(c) 単語辞書を検索することにより、前記単語表記に対応する複数の読み候補と各々が前記複数の読み候補に対応する読みスコアを取得するとともに、ユーザにより前記複数の読み候補のうちの1つの読みが選択されたことをに応答して、該選択された読みを取得する読み生成部と、

(d) 発音辞書を検索することにより、前記選択された読みに対応するベースフォームと発音スコアを取得するベースフォーム生成部と、

(e) 前記発音スコアが所定の閾値を越えている場合には、該取得したベースフォームが登録される音声認識辞書と、

を含み、

(a 2) 前記音声認識エンジンは、ユーザの発声によって入力された第1の音声情報が所定の条件に一致するか否かを判断し、前記音声情報が所定の条件に一致すると判断した場合には、前記表示画面上に認識単語入力フィールドと、読み入力フィールドを有する音声認識ウィザード・パネルを管理する音声認識ウィザードを起動し、

(c 2) 前記読み生成部は、前記音声認識ウィザード・パネルに入力された認識単語と読みに対応して行なった発声に基づく第2の音声情報を取得し、

(d 2) 前記ベースフォーム生成部は、前記第2の音声情報と、前記認識単語と、前記読みとに基づいて、第2のベースフォームを特定し、

(e 2) 前記音声認識辞書は、前記第2のベースフォームを保持することを特徴とする音声認識装置が提供される。

【0 0 2 0】

本発明の他の一態様においては、

表示画面、音声入力装置を備える音声認識装置上で実施される認識単語の登録処理プログラムを格納した記録媒体であって、

該プログラムは、

(a) ユーザによって特定された単語表記を取得することを前記音声認識装置に指示するためのプログラムコードと、

(b) 単語辞書を検索することにより、前記単語表記に対応する読みを取得することを前記音声認識装置に指示するためのプログラムコードと、

(c) 発音辞書を検索することにより、該取得した読みに対応するベースフォームを取得することを前記音声認識装置に指示するためのプログラムコードと、

(d) 該取得したベースフォームを音声認識辞書に登録することを前記音声認識装置に指示するためのプログラムコードと、

を含む記憶媒体が提供される。

【0 0 2 1】

本発明の他の一態様においては、

表示画面、音声入力装置を備える音声認識装置上で実施される認識単語の登録処理プログラムを格納した記録媒体であって、

該プログラムは、

(a) ユーザによって特定された単語表記を取得することを前記音声認識装置に指示するためのプログラムコードと、

(b) 単語辞書を検索することにより、前記単語表記に対応する複数の読み候補と各々が前記複数の読み候補に対応する読みスコアを取得することを前記音声認識装置に指示するためのプログラムコードと、

(c) ユーザにより前記複数の読み候補を表示することを前記音声認識装置に指示するためのプログラムコードと、

(d) ユーザにより前記複数の読み候補のうちの1つの読みが選択されたことをにตอบสนองして、該選択された読みを取得することを前記音声認識装置に指示するためのプログラムコードと、

(e) 発音辞書を検索することにより、前記選択された読みに対応するベースフォームと発音スコアを取得することを前記音声認識装置に指示するためのプロ

グラムコードと、

(f) 前記発音スコアが所定の閾値を越えているか否かを判断することを前記音声認識装置に指示するためのプログラムコードと、

(g) 前記発音スコアが所定の閾値を越えている場合には、該取得したベースフォームを音声認識辞書に登録することを前記音声認識装置に指示するためのプログラムコードと、

を含む記憶媒体が提供される。

【0 0 2 2】

本発明の他の一態様においては、

表示画面、音声入力装置を備える音声認識装置上で実施される音声認識処理プログラムを格納した記録媒体であって、

該プログラムは、

(a) ユーザの発声によって入力された第 1 の音声情報が所定の条件に一致するか否かを判断することを前記音声認識装置に指示するためのプログラムコードと、

(b) 前記音声情報が所定の条件に一致すると判断した場合には、前記表示画面上に認識単語入力フィールドと、読み入力フィールドを有する音声認識ウィザード・パネルを表示することを前記音声認識装置に指示するためのプログラムコードと、

(c) 前記音声認識ウィザード・パネルに入力された認識単語と読みを取得することを前記音声認識装置に指示するためのプログラムコードと、

(d) 前記ユーザが、前記認識単語と前記読みに対応して行なった発声に基づく第 2 の音声情報を取得することを前記音声認識装置に指示するためのプログラムコードと、

(e) 前記第 2 の音声情報と、前記認識単語と、前記読みとに基づいて、ベースフォームを特定することを前記音声認識装置に指示するためのプログラムコードと、

(f) 前記ベースフォームを音声認識辞書に追加することを前記音声認識装置に指示するためのプログラムコードと、

を含む記憶媒体が提供される。

【 0 0 2 3 】

本発明の他の一態様においては、

表示画面、音声入力装置を備える音声認識装置上で実施される音声認識処理プログラムを格納した記録媒体であって、

該プログラムは、

(a) ユーザによって特定された単語表記を取得することを前記音声認識装置に指示するためのプログラムコードと、

(b) 単語辞書を検索することにより、前記単語表記に対応する複数の読み候補と各々が前記複数の読み候補に対応する読みスコアを取得することを前記音声認識装置に指示するためのプログラムコードと、

(c) ユーザに前記複数の読み候補を表示することを前記音声認識装置に指示するためのプログラムコードと、

(d) ユーザにより前記複数の読み候補のうちの 1 つの読みが選択されたことをに応答して、該選択された読みを取得することを前記音声認識装置に指示するためのプログラムコードと、

(e) 発音辞書を検索することにより、前記選択された読みに対応するベースフォームと発音スコアを取得することを前記音声認識装置に指示するためのプログラムコードと、

(f) 前記発音スコアが所定の閾値を越えているか否かを判断することを前記音声認識装置に指示するためのプログラムコードと、

(g) 前記発音スコアが所定の閾値を越えている場合には、該取得したベースフォームを音声認識辞書に登録することを前記音声認識装置に指示するためのプログラムコードと、

(h) ユーザの発声によって入力された第 1 の音声情報が所定の条件に一致するか否かを判断することを前記音声認識装置に指示するためのプログラムコードと、

(i) 前記音声情報が所定の条件に一致すると判断した場合には、前記表示画面上に認識単語入力フィールドと、読み入力フィールドを有する音声認識ウィザ

ード・パネルを表示することを前記音声認識装置に指示するためのプログラムコードと、

(j) 前記音声認識ウィザード・パネルに入力された認識単語と読みを取得することを前記音声認識装置に指示するためのプログラムコードと、

(k) 前記ユーザが、前記認識単語と前記読みに対応して行なった発声に基づく第2の音声情報を取得することを前記音声認識装置に指示するためのプログラムコードと、

(l) 前記第2の音声情報と、前記認識単語と、前記読みとに基づいて、第2のベースフォームを特定することを前記音声認識装置に指示するためのプログラムコードと、

(m) 前記第2のベースフォームを音声認識辞書に追加することを前記音声認識装置に指示するためのプログラムコードと、

を含む記憶媒体が提供される。

【0024】

【実施例】

A. ハードウェア構成

図1、2は、本発明の好適な実施例における音声認識手順の概要を説明する概念図である。この例においては、まず、Text-To-Speechなどで使われている手法で単語表記401からよみ群407と発音表記409のセットをスコア付きで生成する。生成されたよみ群からユーザが正しいよみを選択する。

【0025】

そして、表記-よみからベースフォーム群をスコア付きで生成し、スコアが基準値を超えるものを声の録音なしで音声認識辞書231に登録する。

【0026】

次に、実際の音声認識処理において、認識エラーがN回を超えた場合等一定のエラーが検出された場合には、ユーザの声を録音し、実際に発声したベースフォームを取得し、音声認識辞書に再登録する。再登録されたベースフォームがその

後の音声認識処理で利用されるようになる。

【0027】

#### B. ハードウェア構成

図3は、本発明の音声認識システム100を実施するためのハードウェア構成の概観図である。音声認識システム100は、中央処理装置（CPU）1とメモリ4とを含んでいる。CPU1とメモリ4は、バス2等を介して、補助記憶装置としてのハードディスク装置13、31とを接続してある。フロッピーディスク装置（またはMO28、CD-ROM26、29等の媒体駆動装置26、28、29、30）20はフロッピーディスクコントローラ（またはIDEコントローラ25、SCSIコントローラ27等の各種コントローラ）19を介してバス2へ接続されている。

【0028】

フロッピーディスク装置（またはMO、CD-ROM等の媒体駆動装置26、28、29、30）20には、フロッピーディスク（またはMO、CD-ROM等の媒体）が挿入され、このフロッピーディスク等やハードディスク装置13、ROM14等の記憶媒体には、オペレーティング・システムと協働してCPU等に命令を与え、本発明を実施するためのコンピュータ・プログラムのコードを記録することができ、メモリ4にロードされることによって実行される。このコンピュータ・プログラムのコードは圧縮し、または、複数に分割して、複数の媒体にまたがって記録することもできる。

【0029】

音声認識システム100は更に、ユーザ・インターフェース・ハードウェアを備えたシステムとすることができ、ユーザ・インターフェース・ハードウェアとしては、例えば、画面位置情報を入力するためのポインティング・デバイス（マウス、ジョイスティック、トラックボール等）7、キー入力をサポートするキーボード6や、イメージデータをユーザに提示するためのディスプレイ11、12がある。また、スピーカ23は、オーディオ・コントローラ21から音声信号を、アンプ22を介して受領し、音声として出力する。

【 0 0 3 0 】

この音声認識システム 1 0 0 は、シリアルポート 1 5 およびモデムまたは、トークンリング等の通信アダプタ 1 8 等を介して、他のコンピュータ等と通信を行うことができる。

【 0 0 3 1 】

本発明は、通常のパーソナルコンピュータ（P C）やワークステーション、テレビや F A X 等の各種家電製品に組み込まれたコンピュータ、車両や航空機等に搭載されるコンピュータ（カーナビゲーション・システム等）又はこれらの組合せによって実施可能である。ただし、これらの構成要素は例示であり、その全ての構成要素が本発明の必須の構成要素となるわけではない。特に、本発明は、音声認識辞書のカスタマイズに関するものであるため、パラレルポート 1 6、シリアルポート 1 5 等の構成要素は本発明の一態様においては必須のものではない。

【 0 0 3 2 】

音声認識システム 1 0 0 のオペレーティング・システムとしては、W i n d o w s N T（マイクロソフトの商標）、W i n d o w s 9 x（マイクロソフトの商標）、W i n d o w s 3 . x（マイクロソフトの商標）、O S / 2（I B M の商標）、M a c O S（A p p l e 社の商標）、L i n u x（L i n u s T o r v l d s の商標）、A I X（I B M の商標）上の X - W I N D O W システム（M I T の商標）などの、標準で G U I マルチウインドウ環境をサポートするものや、P C - D O S（I B M の商標）、M S - D O S（マイクロソフトの商標）などのキャラクタ・ベース環境のもの、さらには O S / O p e n（I B M の商標）、V x W o r k s（W i n d R i v e r S y s t e m s , I n c . の商標）等のリアルタイム O S、J a v a O S 等のネットワークコンピュータに組み込まれた O S でも実現可能であり、特定のオペレーティング・システム環境に限定されるものではない。

【 0 0 3 3 】

C. システム構成



図4は、本発明の好適な実施例における音声認識システムの処理要素を示す機能ブロック図である。

【0034】

本発明の好適な実施例における音声認識システムは、認識単語登録部201、音声認識エンジン203、読み生成部205、ベースフォーム生成部207、音声入力部209、音声認識ウィザード211、音声認識辞書231、単語辞書233及び、発音辞書235を備えている。

【0035】

認識単語登録部201は、本発明の単語登録処理のユーザ入力や各種情報の表示をオペレーティング・システム等と共同して制御する。また、読み生成部等、他の処理要素の制御を行なう。

【0036】

音声認識エンジン203は、音声認識辞書を用いて、実際に入力された音声情報を解析し、テキストデータを出力する。また、所定の音声認識エラーを検出し、音声認識ウィザード211を起動する。

【0037】

読み生成部205は、認識単語登録部より渡された、文字表記データを受領し、単語辞書233を用いて、その読みとスコアを探索する。また、所定のロジックに基づいて、読みスコアの補正を行なう。この読みスコアは、発生確率等の統計情報、経験値等に基づいて設定することができる。

【0038】

図5は、本発明の好適な実施例における単語辞書233の概念図である。図に示すように、本発明の好適な実施例における単語辞書233には、単語表記301、読み303、発音表記305、読みスコア307の情報が管理されている。

【0039】

ベースフォーム生成部207は、入力された単語表記及び読み情報を基に、発音辞書を検索し、対応するベースフォームと発音スコアを出力する。また、所定のロジックに基づいて、発音スコアの補正を行なう。この発音スコアは、発生確率等の統計情報、経験値等に基づいて設定することができる。また、単語表記に

対するベースフォームのスコアは、読みスコア×発音スコア等、読みスコアと発音スコアに基づいた関数値を設定することができる。

【0 0 4 0】

図 6 は、本発明の好適な実施例における発音辞書 2 3 5 の概念図である。図に示すように、本発明の好適な実施例における発音辞書 2 3 5 には、発音表記 3 1 1、ベースフォーム 3 1 3、発音スコア 3 1 5 の情報が管理されている。

【0 0 4 1】

音声入力部 2 0 9 は、ユーザが発声した音声情報をシステムに取り込む。

【0 0 4 2】

音声認識ウィザード 2 1 1 は、ユーザの発声によって入力された音声情報を基に、単語表記に対応するベースフォームを決定する。

【0 0 4 3】

以上図 2 に示す各機能ブロックを説明したが、これらの機能ブロックは、論理的機能ブロックであり、各々 1 つのまとまりを持ったハードウェアやソフトウェアによって実現されることを意味するのではなく、複合し、または共通したハードウェアやソフトウェアによって実現可能である。

【0 0 4 4】

#### D. 単語登録手順

次に、図 7 を用いて、本発明の好適な実施例における認識単語登録手順を説明する。まず、図 8 の本発明の好適な実施例における認識単語登録部のユーザインタフェースに示されるように、認識単語登録部 2 0 1 は、ユーザーが登録単語フィールド 4 0 1 に入力した登録したい単語の表記を取得する（ステップ 4 5 3）。

【0 0 4 5】

そして、入力された単語表記が読み生成部 2 0 5 に渡され、読み生成部 2 0 5 により、単語辞書 2 3 3 の検索が行われる（ステップ 4 5 5）。読み生成部 2 0 5 は、その単語表記に対応する読み 3 0 3、発音表記 3 0 5、読みスコア 3 0 7 を取得する（ステップ 4 5 7）。なお、一定の条件に適合する場合には、取得さ

れたスコア 307 を読み生成部 205 のロジックにより補正することもできる。

#### 【0046】

例えば、表記がかなのみで構成され、確実に長音化されないケースでは、スコアを 1 に補正する。また、表記にマッチするものが辞書に存在する場合であっても、その辞書の精度が音声認識的には高くない場合（例えばかな漢字変換用の辞書を流用した場合等）、読みが長音化される可能性があればスコアを下げる補正を行なう。

#### 【0047】

この読み、発音表記、読みスコアは、読み生成部 205 から認識単語登録部 201 に返され、読み候補が生成されたか否かが判断される（ステップ 459）。読み候補が生成されなかった場合には、認識単語登録部 201 は、ユーザに読みの入力を指示するメッセージを出力し、入力された読み 407、発音表記 409 を取得する（ステップ 461）。図 9 は、ユーザがフィールド 407、409 にキーボード入力した状態を示している。

#### 【0048】

読み候補が生成された場合には、スコアが最高の読みを読みフィールド 409 に、その発音表記を発音表記フィールド 409 に表示する。また、図 10 に示すように、プルダウンメニューによりスコア順に並んだ読みをユーザが選択できるようにする（ステップ 463）。

#### 【0049】

ユーザは、現在読みフィールド 407 に表示されている読み、発音表記フィールド 409 に表示されている発音表記を確認し、正しくなければ他の候補を選択するか、キーボード入力により修正する。

#### 【0050】

認識単語登録パネル 400 において、正しい読み 407、発音表記 409 が入力された状態において、ユーザが「単語を登録」ボタン 403 を押すと、認識単語登録部 201 が、単語表記 401、読み 407、発音表記 409 のセットを取得し、これをベースフォーム生成部 207 に渡す。なお、入力された読みが不適当な場合（「あゝ」など）には、ユーザに正しい入力を促すメッセージを出力す

る。

#### 【0051】

ベースフォーム生成部207は、この情報を基に発音辞書235を探索し、対応するベースフォームと発音スコアを取得する。本発明の好適な実施例においては、単語表記、発音表記を用いて、発音スコアが補正される。

#### 【0052】

そして、取得したベースフォームのうち、スコアが所定の閾値を越えてたものは、表記-よみ-ベースフォームのセットで音声認識辞書に登録する（ステップ467）。このような手順により、ユーザによる実際の発声作業を伴わず、新しい単語を音声認識辞書に登録することができる。

#### 【0053】

### E. 音声認識手順

次に、図11を用いて、本発明の好適な実施例における音声認識手順を説明する。まず、ユーザは、音声認識エンジン203を起動し、音声入力部209よりユーザの発声により生成した音声情報を入力する。音声認識エンジン203は、音声認識辞書231を使用し、公知の音声認識手法を用いて、入力された音声情報を認識し、テキストデータを出力する。

#### 【0054】

この音声認識処理において、音声認識エンジン203が、所定のエラーを検出した場合には（ステップ483）、そのエラーが所定の条件に一致するか否かを判断する（ステップ487）。

例えば、

1. 認識時のスコアが低かったためにリジェクトされた（たとえば3回）
2. 同じ単語に対する修正が複数回（たとえば2回）行われた
3. 誤認識が複数回（たとえば3回）続いた

といった条件を考えられる。

係る場合、エラーが認識された単語毎エラーカウンターを持たせ、エラーをカウントする等の方法で実現することができる。

## 【0 0 5 5】

そして、この条件に一致したと判断された場合には、音声認識エンジン 2 0 3 は、音声認識ウィザード 2 1 1 を起動し、音声認識ウィザードが終了するまで処理を中断する（ステップ 4 8 9）。

## 【0 0 5 6】

音声認識ウィザード 2 1 1 は、図 1 3、1 4 で説明した従来の認識単語登録方法と同様の手順で認識単語に対応するベースフォームを選択することができる。図 1 2 は、本発明の好適な実施例における音声認識ウィザード 2 1 1 のユーザインタフェースを示す概念図である。図に示すように、ユーザは、音声認識ウィザード 2 1 1 の読みフィールド 6 0 1、発音表記フィールド 2 1 3 に読みと発音表記を入力し、録音ボタン（図示せず：一旦録音ボタンが押されると、録音ボタンが消え、「録音中」を示すメッセージが表示される）を押して単語の読み上げを行なうことにより、音声情報を取得し、ユーザーの声のゆれをベースフォーム上に反映させることにより、音声認識辞書への登録を行ない、その後の音声認識で利用できるようにする。

## 【0 0 5 7】

このように本発明の方法でも最終的に声を録音しなければならない場合は存在する。これは方言があったり、標準的な発音からかけ離れすぎた人の場合等が考えられる。しかし、ほとんどの場合は、実際にユーザは、単語の登録際し、音声による登録なしに作業を終了することができ、従来の方法では必ず声の録音が必要だったことに比較して、大幅にユーザーの登録時の作業量が削減できる。また認識エラーがある場合はユーザーに声を録音するように自動的に促されるため、認識率の低下を回避することが可能である。

## 【0 0 5 8】

なお上記記述では理解を深めるため日本語を例にとったが、本発明は日本語だけに限定された発明ではない。基本的にはすべての言語で利用可能である。これは、他の言語専用の適切な読み生成部、ベースフォーム生成部、音声認識エンジン等を用意することにより、本発明をそのまま適用することができるからである。

【0059】

この場合、各国で標準化された読みの表現方法を採用する必要がある。例えば日本語においては「ひらがな」、中国ではP i n Y i n、台湾ではJ u Y i n、韓国ではハングル等が考えられる。欧米の言語については、英語を除き、読みの表現方法は一般的ではないため、かかる言語に適用する場合には、読み表現のゆらぎを吸収して設計する必要がある。

【0060】

【発明の効果】

以上説明したように、本発明によれば、音声認識の精度を維持しつつ、ユーザーの音声による録音の作業を可能な限り低減させることができる。

【0061】

【図面の簡単な説明】

【図1】 本発明の好適な実施例における認識単語登録手順の概要を説明する概念図である。

【図2】 本発明の好適な実施例における認識単語登録手順の概要を説明する概念図である。

【図3】 本発明の音声認識装置のハードウェア構成の一実施例を示すブロック図である。

【図4】 本発明の好適な実施例における音声認識システムの処理要素のブロック図である。

【図5】 本発明の好適な実施例における単語辞書の概念図である。

【図6】 本発明の好適な実施例における発音辞書の概念図である。

【図7】 本発明の好適な実施例における認識単語登録手順を示すフローチャートである。

【図8】 本発明の好適な実施例における認識単語登録部のユーザインタフェースを示す概念図である。

【図9】 本発明の好適な実施例における認識単語登録部のユーザインタフェースを示す概念図である。

【図10】 本発明の好適な実施例における認識単語登録部のユーザインタ

フェースを示す概念図である。

【図 1 1】 本発明の好適な実施例における音声認識手順を示すフローチャートである。

【図 1 2】 本発明の好適な実施例における音声認識ウィザードのユーザインタフェースを示す概念図である。

【図 1 3】 従来の認識単語登録手順の概要を説明する概念図である。

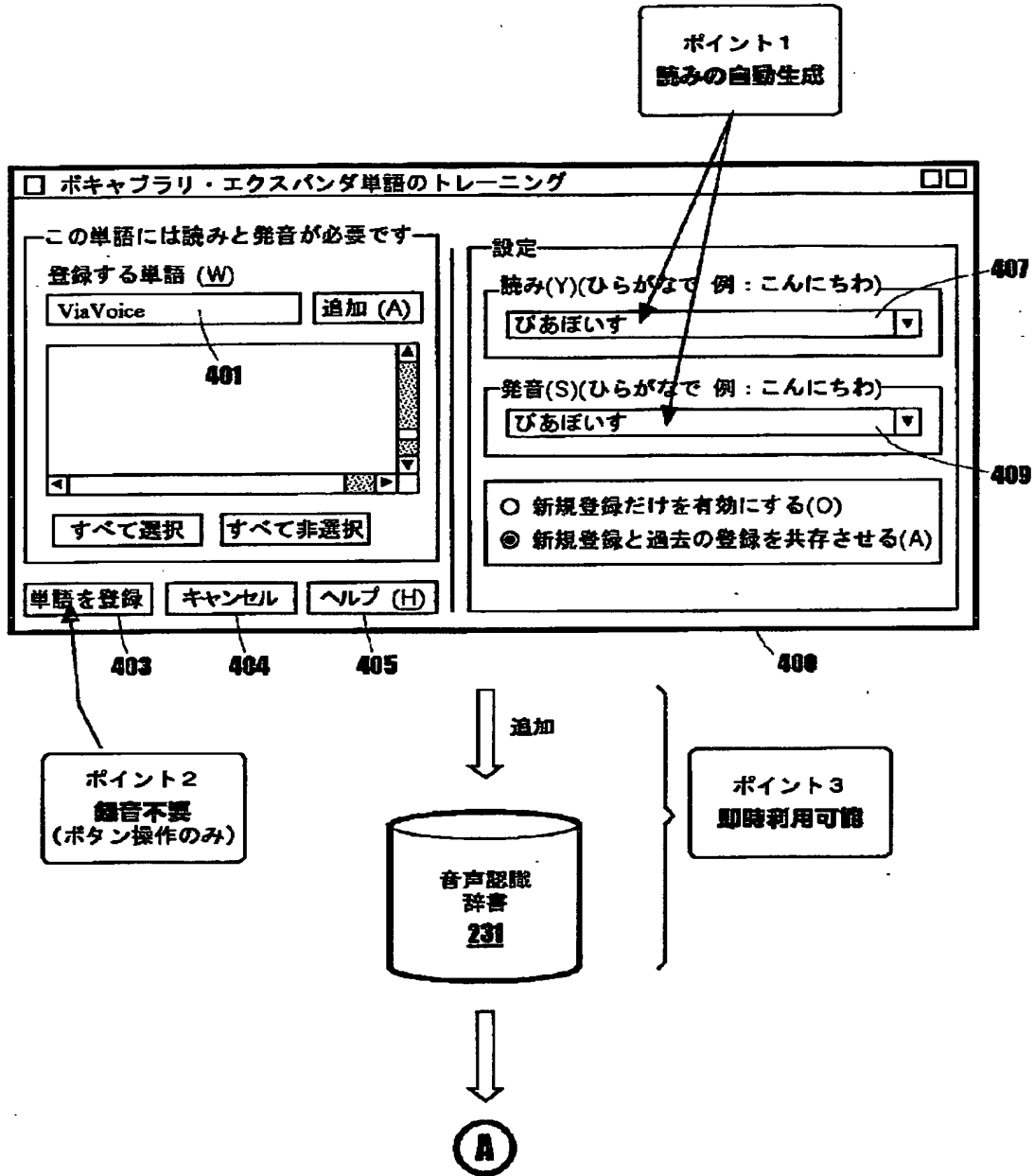
【図 1 4】 従来の認識単語登録手順の概要を説明する概念図である。

【符号の説明】

- 1 0 0 音声認識システム
- 2 0 1 認識単語登録部
- 2 0 3 音声認識エンジン
- 2 0 5 読み生成部
- 2 0 7 ベースフォーム生成部
- 2 0 9 音声発生部
- 2 1 1 音声認識ウィザード
- 2 3 1 音声認識辞書
- 2 3 3 単語辞書
- 2 3 5 発音辞書

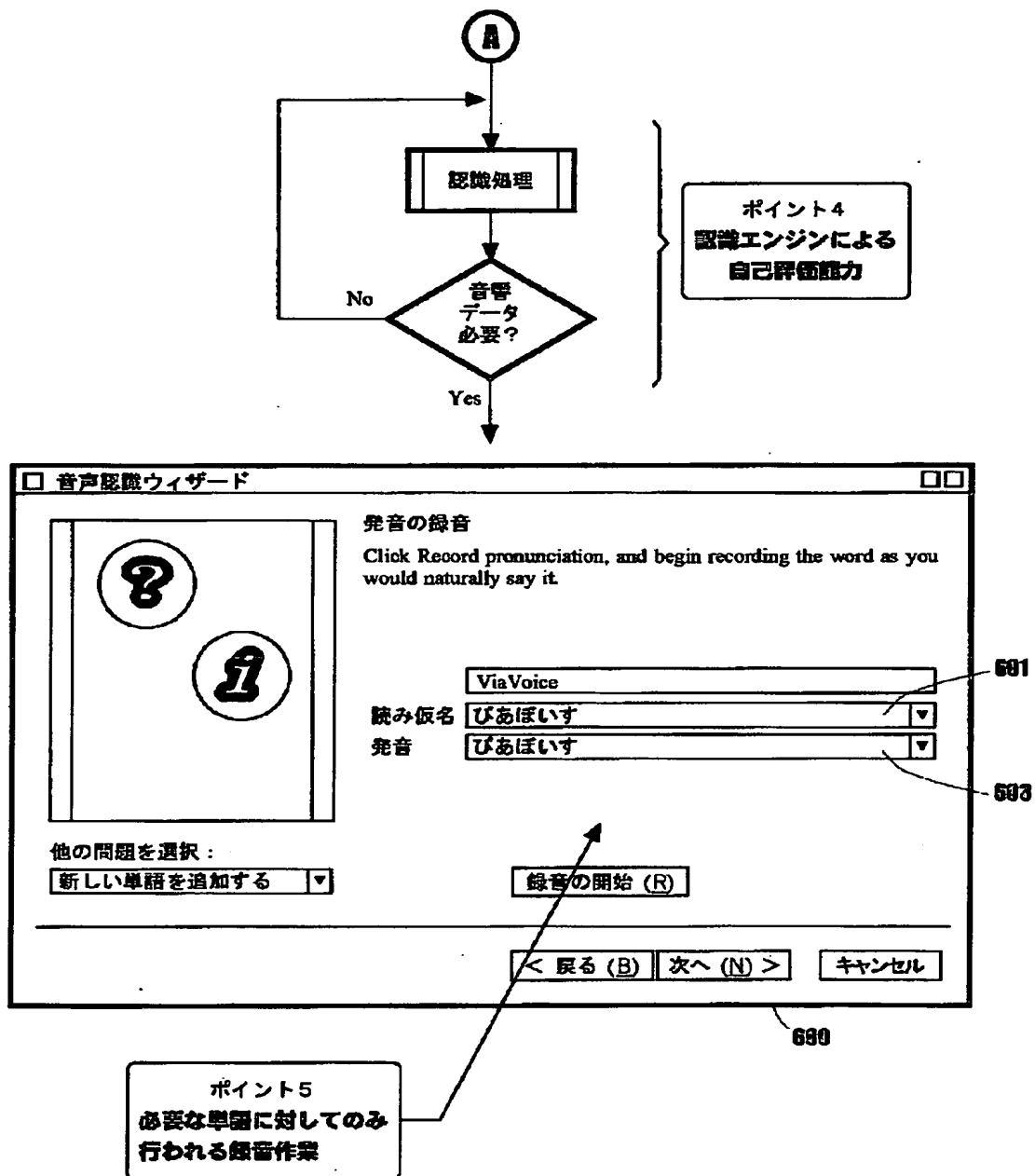
【書類名】 図面

【図 1】

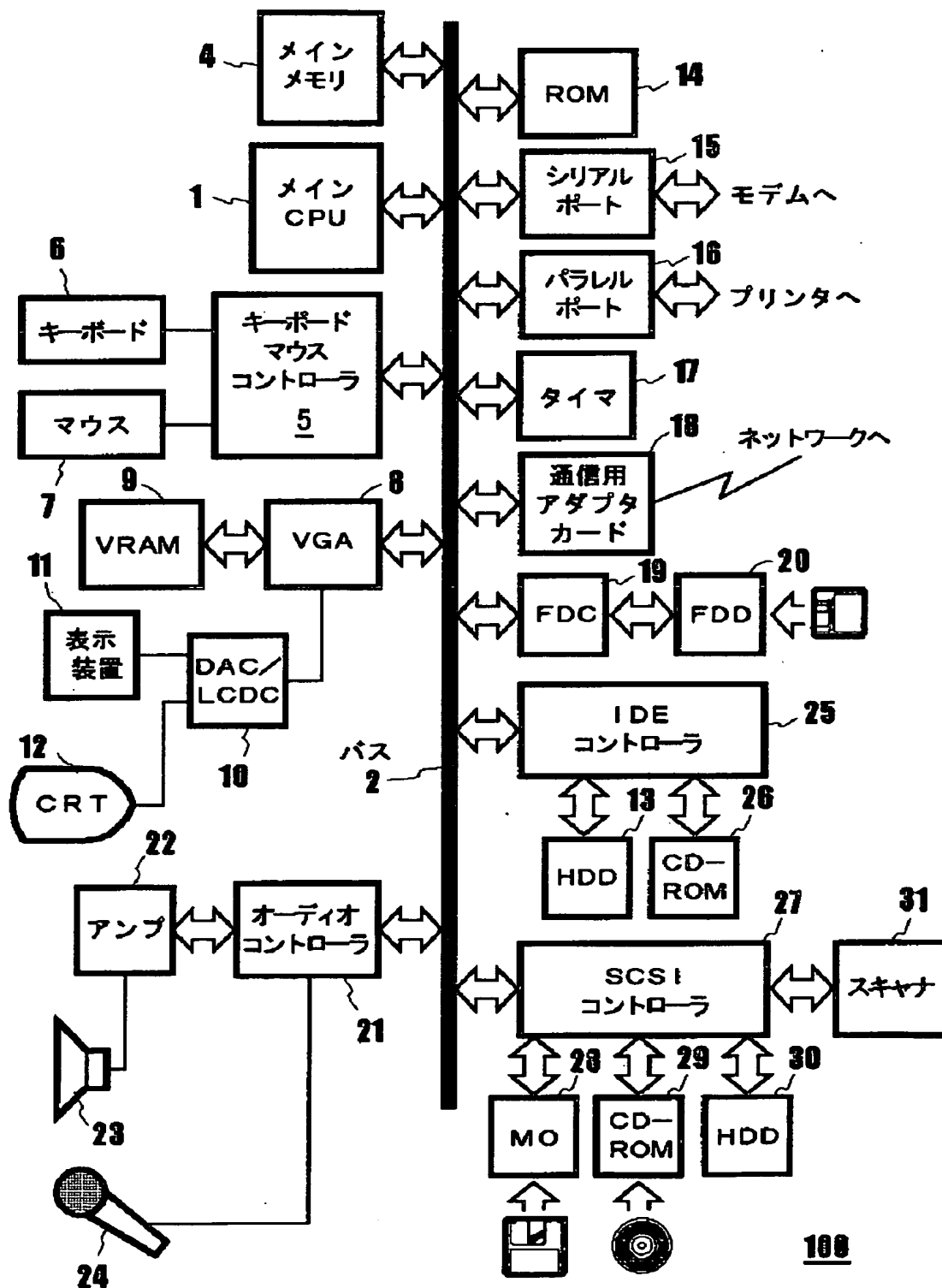




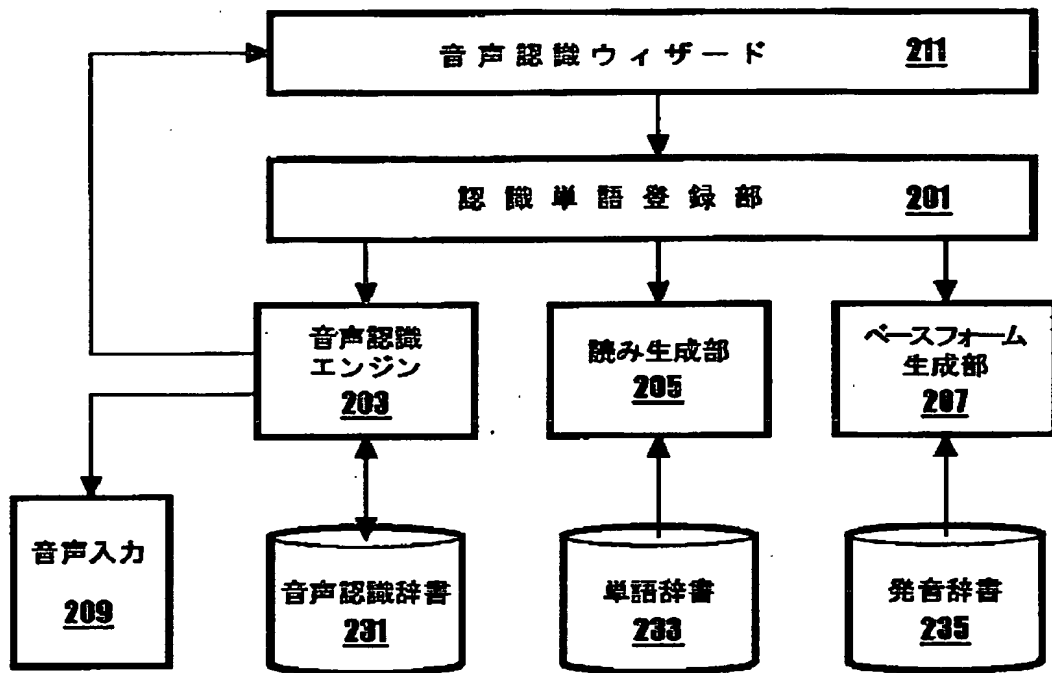
【図 2】



【図 3】



【図4】



【図5】

単語表記 301		
田 原		
読み 303	発音表記 305	読みスコア 307
たはら	tahara	0.83
たわら	tawara	0.56
たばら	tabara	0.45
たばる	tabaru	0.20
だはら	dahara	0.02

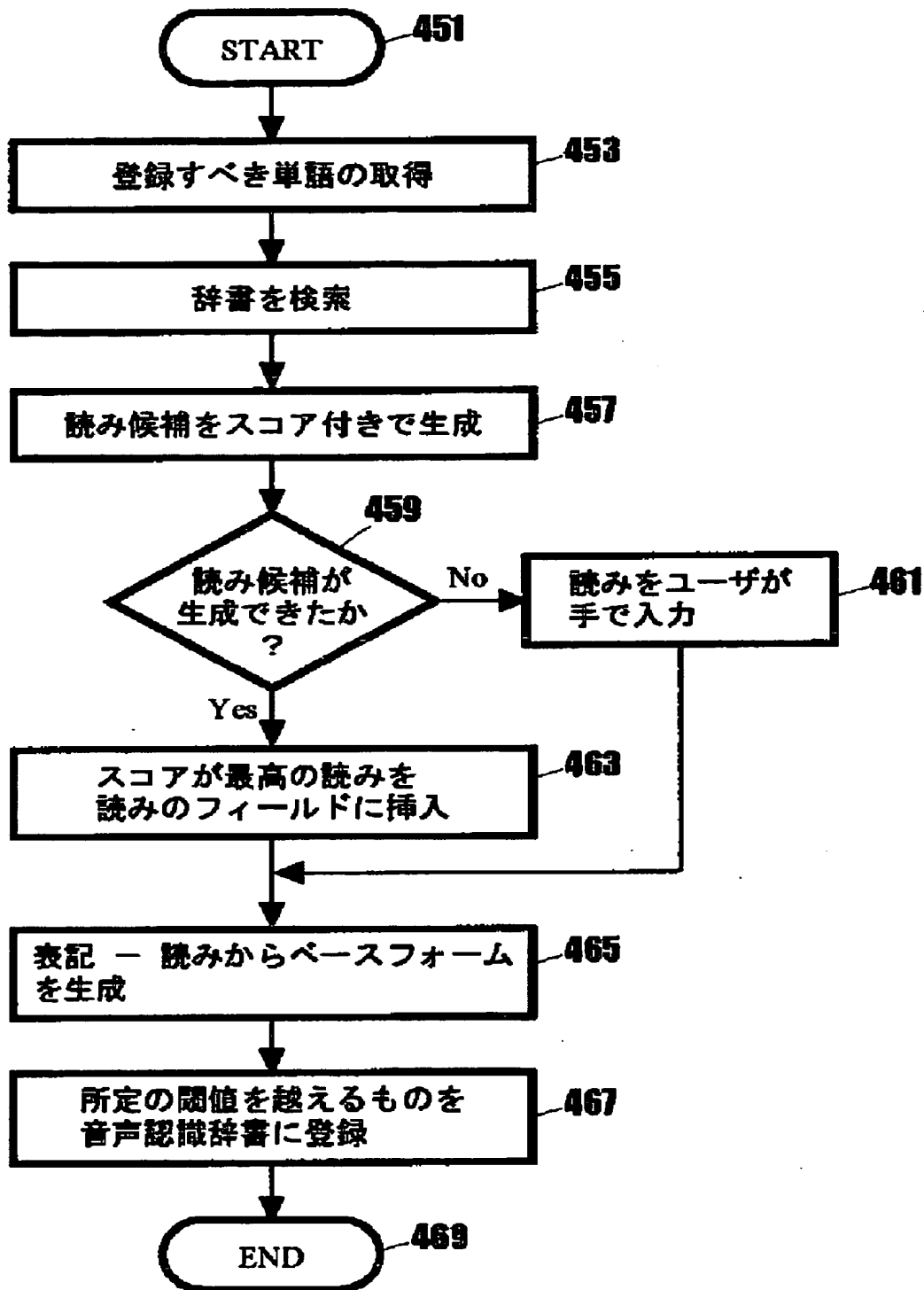
308

【図 6】

発 音 表 記 <u>311</u>	
こ う し	
ベースフォーム <u>313</u>	発音スコア <u>315</u>
ベースフォーム 1	0.74
ベースフォーム 2	0.58
ベースフォーム 3	0.37

318

【図 7】



【図 8】

□ ポキャブラリ・マネージャ 単語のトレーニング □□

登録する単語を選択してください ...

登録する単語 (W) ViaVoice 追加 (L) 481

☒ ViaVoice 482

すべて選択 (E)
すべて非選択 (C)

単語を登録 (R)
キャンセル
ヘルプ (H)
483      484      485

設定

読み(Y)(ひらがなで 例: こんにちは) 487

ぴあほいす ▼

発音(S)(ひらがなで 例: こんにちは)

ぴあほいす ▼

☐ 新規登録だけを有効にする(O)
 ☒ 新規登録と過去の登録を共存させる(A) 489

☐ 発音が不明な単語を優先して表示 (B)

489

【図 9】

□ ポキャブラリ・マネージャ 単語のトレーニング □□

登録する単語を選択してください ...

登録する単語 (W) おかあさんを 追加 (L) 481

すべて選択 (E)
すべて非選択 (C)

単語を登録 (R)
キャンセル
ヘルプ (H)
483      484      485

設定

読み(Y)(ひらがなで 例: こんにちは) 487

おかーさんを ▼

発音(S)(ひらがなで 例: こんにちは)

おかーさんお ▼

☐ 新規登録だけを有効にする(O)
 ☒ 新規登録と過去の登録を共存させる(A) 489

☐ 発音が不明な単語を優先して表示 (B)

489

489

【図 10】

□ ポキャブラリ・マネージャ 単語のトレーニング □□

登録する単語を選択してください ...

登録する単語 (W) 401

田原 追加 (L)

☒ 田原 402

すべて選択 (E) すべて非選択 (C)

単語を登録 (R) 403 キャンセル 404 ヘルプ (H) 405

設定

読み(Y)(ひらがなで 例: こんにちは) 407

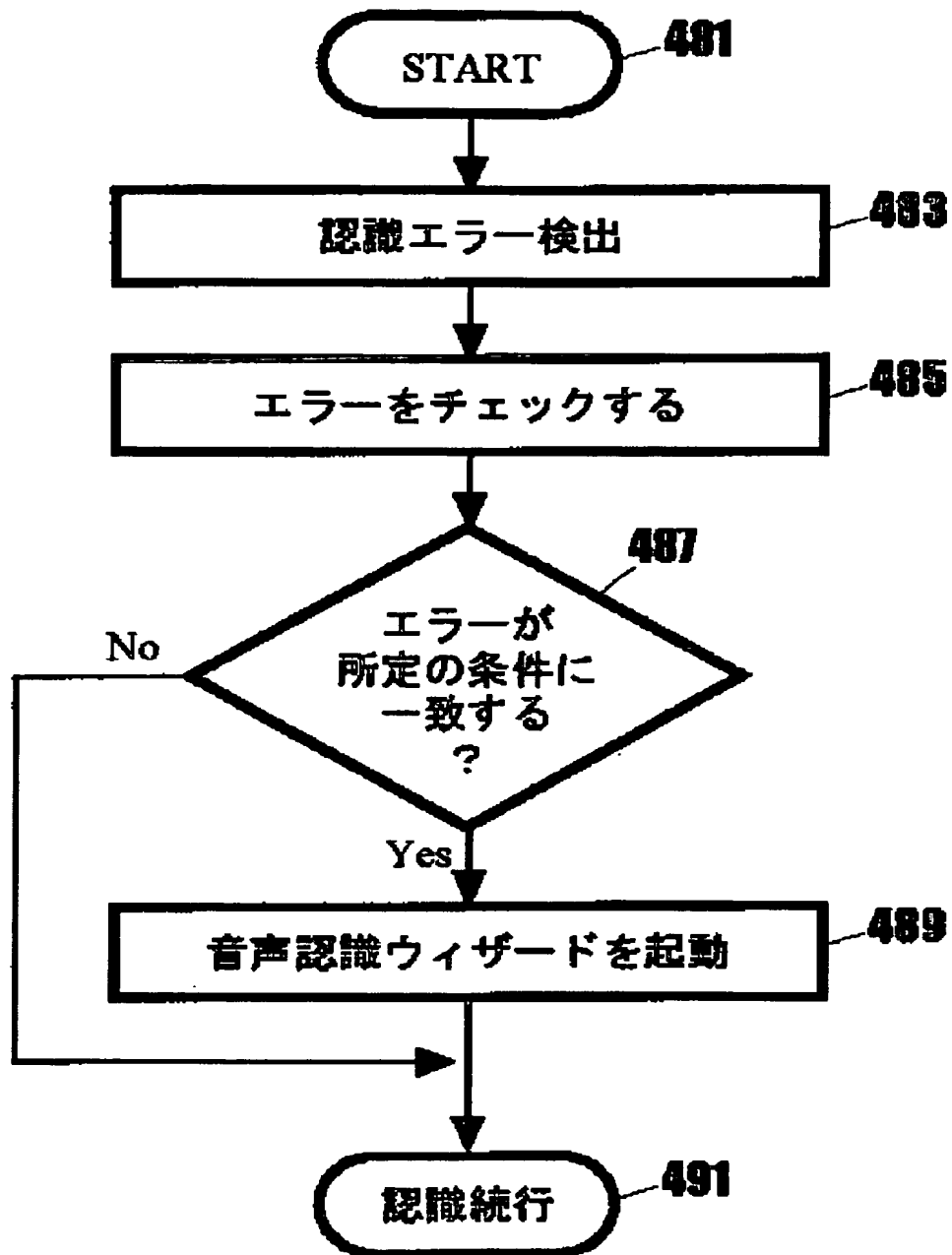
たはら  
たはら  
たわら  
たばら  
たばる

☐ 新規登録だけを有効にする(O)

☒ 新規登録と過去の登録を共存させる(A)

☐ 発音が不明な単語を優先して表示 (B) 400

【図 11】





【図 12】

音声認識ウィザード

Teach ViaVoice your pronunciation of a word  
Click Record pronunciation, and begin recording the word as you would naturally say it.

他の問題を選択:  
コマンドの単語を登録する ▾

読み仮名  ▾ 601  
発音  ▾ 603

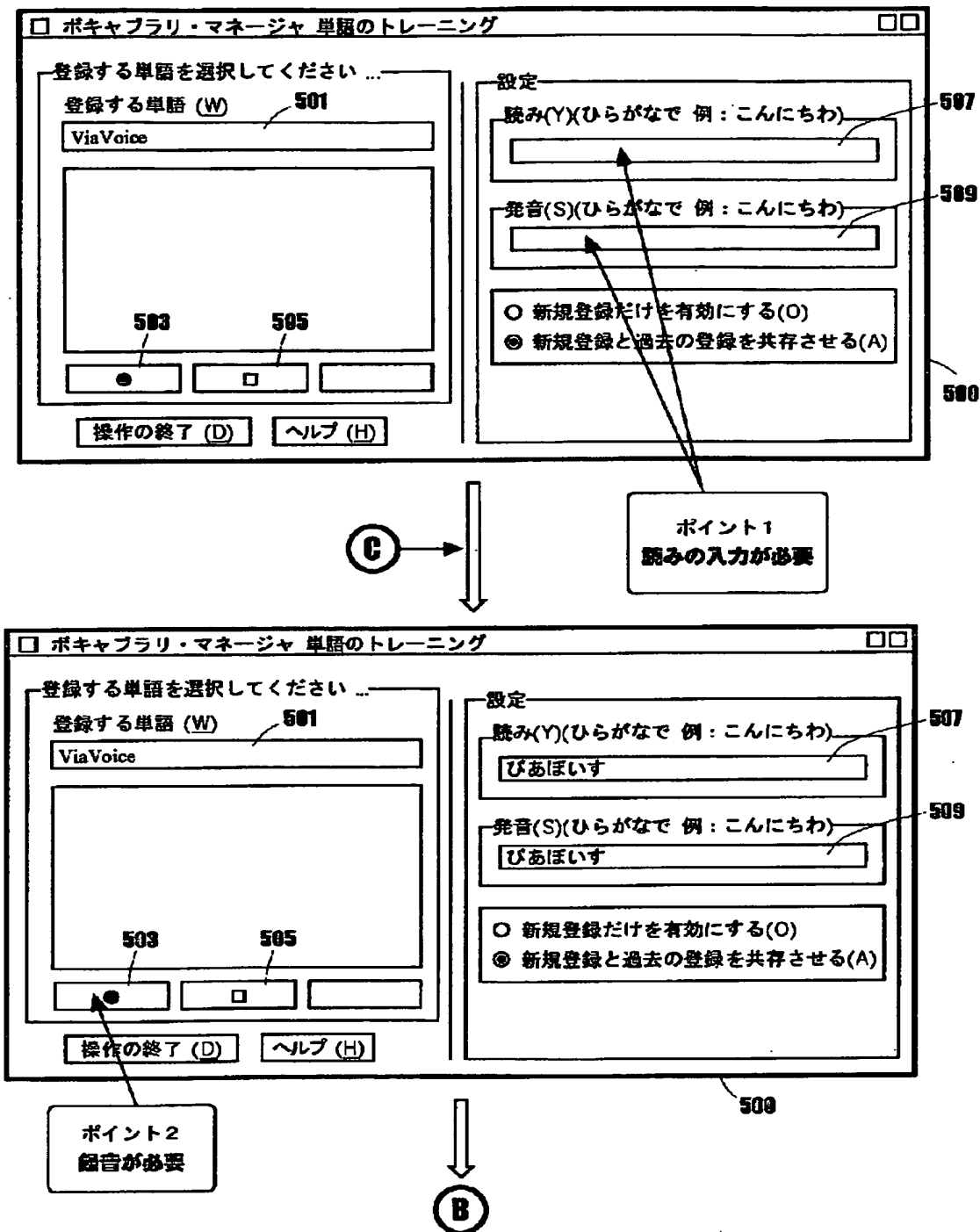
録音中 ....

コマンドを発生してください。  
単語の登録

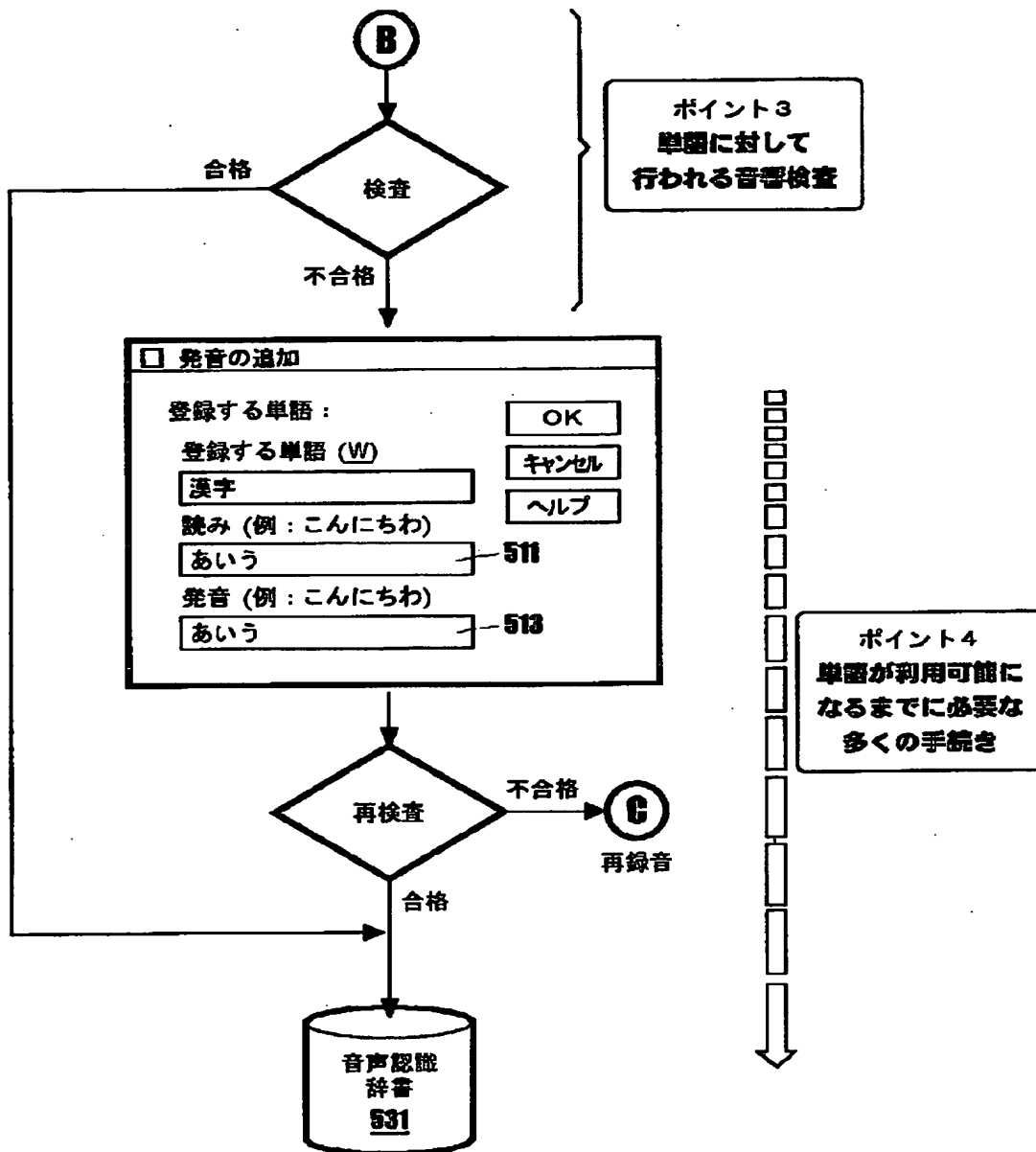
< 戻る (B) 次へ (N) > キャンセル

600

【図 1 3】



【図 14】



【書類名】 要約書

【要約】

【課題】 音声認識の精度を維持しつつ、ユーザーの音声による録音の作業を可能な限り低減させる。

【解決手段】

単語表記から読み群をスコア付きで生成する。生成された読み群からユーザーが正しいよみを選択する。選択された読みからと単語表記-よみからベースフォーム群をスコア付きで生成し、スコアが基準値を超えるものを声の録音なしで登録する。ユーザが音声認識処理を行なうに際し、認識エラーがN回を超えた場合等所定のエラー状態を検出した場合には、ユーザに単語表記、読み、発音表記を登録し、対応する音声を発声することを促すパネルを出力する。ユーザーの実際の発声にもとづいてベースフォームを取得し、音声認識辞書に再登録する。

【選択図】 図4

認定・付加情報

特許出願の番号	平成11年 特許願 第257429号
受付番号	59900885023
書類名	特許願
担当官	第七担当上席 0096
作成日	平成11年 9月14日

<認定情報・付加情報>

【提出日】	平成11年 9月10日
-------	-------------

出 願 人 履 歴 情 報

識別番号 [390009531]

1. 変更年月日 1990年10月24日

[変更理由] 新規登録

住 所 アメリカ合衆国10504、ニューヨーク州 アーモンク (番地なし)

氏 名 インターナショナル・ビジネス・マシーンズ・コーポレイション